PATENT COOPERATION TREATY

●10/019615

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

PCT

WRITTEN OPINION

(PCT Rule 66)

To: RONALD E. LARSON MCANDREWS HELD & MALLOY, LTD. 500 W. MADISON, 34TH FLOOR CHICAGO IL 60661	Date of Mailing (day/month/year) **2 7 AUG 2001**

Applicant's or agent's file reference 12785W001	REPLY DUE within TWO months from the above date of mailing

International application No. PCT/US00/18104	International filing date (day/month/year) 30 JUNE 2000	Priority date (day/month/year) 02 JULY 1999

International Patent Classification (IPC) or both national classification and IPC
IPC(7): H04B 1/66 and US Cl.: 375/240

Applicant
TELLABS OPERATIONS, INC.

1. This written opinion is the _first_ (first, etc.) drawn by this International Preliminary Examining Authority.

2. This opinion contains indications relating to the following items:

    I   [X] Basis of the opinion

    II [ ] Priority

    III [ ] Non-establishment of opinion with regard to novelty, inventive step or industrial applicability

    IV [ ] Lack of unity of invention

    V   [X] Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability;
             citations and explanations supporting such statement

    VI [ ] Certain documents cited

    VII [ ] Certain defects in the international application

    VIII [X] Certain observations on the international application

3. The applicant is hereby invited to reply to this opinion.

    When?   See the time limit indicated above. ~~The applicant may, before the expiration of that time limit, request this Authority to grant an extension., see Rule 66.2(d).~~

    How?   By submitting a written reply, accompanied, where appropriate, by amendments, according to Rule 66.3.
                For the form and the language of the amendments, see Rules 66.8 and 66.9.

    Also   For an additional opportunity to submit amendments, see Rule 66.4.
                For the examiner's obligation to consider amendments and/or arguments, see Rule 66.4 *bis*.
                For an informal communication with the examiner, see Rule 66.6.

    If no reply is filed, the international preliminary examination report will be established on the basis of this opinion.

4. The final date by which the international preliminary
    examination report must be established according to Rule 69.2 is: _02 NOVEMBER 2001_

Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer MOHAMMAD GHAYOUR Telephone No. (703) 306-3034

Form PCT/IPEA/408 (cover sheet) (July 1998)★

I. Basis of the opinion

1. With regard to the elements of the international application:*

[X] the international application as originally filed

[X] the description:

pages _____ 1-31 _____ , as originally filed

pages _____ NONE _____ , filed with the demand

pages _____ NONE _____ , filed with the letter of _____

[X] the claims:

pages _____ 32-41 _____ , as originally filed

pages _____ NONE _____ , as amended (together with any statement) under Article 19

pages _____ NONE _____ , filed with the demand

pages _____ NONE _____ , filed with the letter of _____

[X] the drawings:

pages _____ 1-8 _____ , as originally filed

pages _____ NONE _____ , filed with the demand

pages _____ NONE _____ , filed with the letter of _____

[X] the sequence listing part of the description:

pages _____ NONE _____ , as originally filed

pages _____ NONE _____ , filed with the demand

pages _____ NONE _____ , filed with the letter of _____

2. With regard to the language, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.
These elements were available or furnished to this Authority in the following language _____ which is:

[ ] the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).

[ ] the language of publication of the international application (under Rule 48.3(b)).

[ ] the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the written opinion was drawn on the basis of the sequence listing:

[ ] contained in the international application in printed form.

[ ] filed together with the international application in computer readable form.

[ ] furnished subsequently to this Authority in written form.

[ ] furnished subsequently to this Authority in computer readable form.

[ ] The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

[ ] The statement that the information recorded in computer readable form is identical to the writen sequence listing has been furnished.

4. [X] The amendments have resulted in the cancellation of:

[X] the description, pages _____ NONE _____

[X] the claims, Nos. _____ NONE _____

[X] the drawings, sheets/fig _____ NONE _____

5. [ ] This opinion has been drawn as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).

* *Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this opinion as "originally filed".*

Form PCT/IPEA/408 (Box I) (July 1998)*

V. Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. statement

Novelty (N)	Claims	1-61	YES
	Claims	NONE	NO
Inventive Step (IS)	Claims	1-61	YES
	Claims	NONE	NO
Industrial Applicability (IA)	Claims	1-61	YES
	Claims	NONE	NO

2. citations and explanations

Claims 1-61 meet the criteria set out in PCT Article 33(2)-(4), because the prior art does not teach or fairly suggest a method and apparatus for reducing echo in a near end digital signal in accordance to partially decoded near end and far end signals and based on adjusting at least a first parameter in the near end digital signal.

------------------ NEW CITATIONS ------------------
NONE

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

Claims 26-31 are objected to under PCT Rule 66.2(a)(v) as lacking clarity under PCT Article 6 because the claims 26-31 are indefinite for the following reason(s): as to claim 26, it an incomplete claim because line 6 of the claim is not continued (i.e "comprising:" what?). as to claim 27, it is not written as an independent claim, nor it does depend on any previous claim. as to claims 28-31, they depend on an indefinite parent claim (i.e. claim 26).

Supplemental Box
(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: Boxes I - VIII   Sheet 10

TIME LIMIT:
        The time limit set for response to a Written Opinion may not be extended. 37 CFR 1.484(d). Any response received after the expiration of the time limit set in the Written Opinion will not be considered in preparing the International Preliminary Examination Report.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/18104

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :H04B 1/66

US CL :375/240

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 375/240, 245, 324, 340; 332/123, 124, 127; 704/200, 207, 211, 500.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EAST (near end, far end, compress$3, decod$3)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A,E	US 6,112,177 A (COSATTO et al) 29 AUGUST 2000, figures 3a and 3b.	1, 26, 32, and 57
Y,P	US 6,064,693 A (OLIVER et al) 16 MAY 2000, abstract, figures 1 and 2, col. 1, line 43 to col. 3, line 45.	1-61
Y	US 5,847,303 A (MATSUMOTO) 08 DECEMBER 1998, abstract, figures 8, 10, 11, and 13, col. 1, line 14 to col. 3, line 36.	1-61
Y	US 5,828,995 A (SATYAMURTI et al) 27 OCTOBER 1998, abstract, figures 2, 25, 26.	1-61
Y	US 5,140,543 A (TANAKA) 18 AUGUST 1992, abstract, figure 2.	1-61

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)		
"O"	document referring to an oral disclosure, use, exhibition or other means	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 SEPTEMBER 2000	**25 OCT 2000**

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	MOHAMMAD GHAYOURMANESH
Facsimile No. (703) 305-3230	Telephone No. (703) 306-3934

Form PCT/ISA/210 (second sheet) (July 1998) ★

(54) Title: CODED DOMAIN ECHO CONTROL
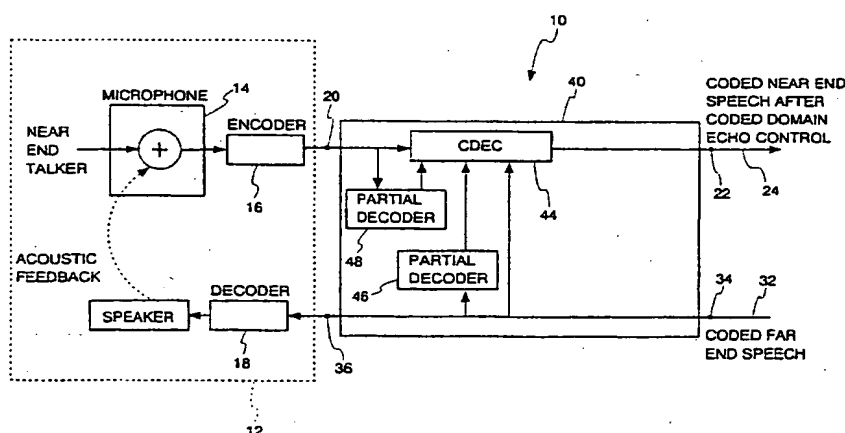
(57) Abstract: A communications system (10) transmits a near end digital signal using a compression code comprising a plurality of parameters including a first parameter. The parameters represent an audio signal comprising a plurality of audio characteristics. The compression code is decodable by a plurality of decoding steps. The system also transmits a far end digital signal using a compression code. A terminal (20) receives the near end digital signal, and a terminal (36) receives the far end digital signal. A processor (40) is responsive to the near end digital signal to read at least the first parameter. The processor generates at least partially decoded near end signals and at least partially decoded far end signals. Based on such signals, the processor adjusts the first parameter and writes the adjusted first parameter into the near end digital signal. Another terminal (22) transmits the adjusted near end digital signal. As a result, the echo in the near end digital signal is reduced.

TITLE OF THE INVENTION

CODED DOMAIN ECHO CONTROL

CROSS-REFERENCE TO RELATED APPLICATIONS

This is a utility application corresponding to provisional application no. 60/142,136 entitled

5   "CODED DOMAIN ENHANCEMENT OF COMPRESSED SPEECH " filed July 2, 1999.

STATEMENT REGARDING FEDERALLY

SPONSORED RESEARCH OR DEVELOPMENT

Not Applicable.

BACKGROUND OF THE INVENTION

10   The present invention relates to coded domain enhancement of compressed

speech and in particular to coded domain echo contol.

This specification will refer to the following references:

[1] GSM 06.10, "Digital cellular telecommunication system (Phase 2); Full rate speech; Part 2: Transcoding", ETS 300 580-2, March 1998, Second Edition.

15   [2] GSM 06.60, "Digital cellular telecommunications system (Phase 2); Enhanced Full Rate (EFR) speech transcoding", June 1998.

[3] GSM 08.62, "Digital cellular telecommunications system (Phase 2+); Inband Tandem Free Operation (TFO) of Speech Codecs", ETSI, March 2000.

[4] J. R. Deller, J. G. Proakis, J. H. L. Hansen, "Discrete-Time Processing of Speech Signals", Chapter
20   7, Prentice-Hall Inc, 1987.

[5] GSM 06.12, "European digital cellular telecommunications system (Phase 2); Comfort noise aspect for full rate speech traffic channels", ETSI, September 1994.

In the GSM digital cellular network, speech transmission between the mobile

25   stations (handsets) and the base station is in compressed or coded form. Speech

coding techniques such as the GSM FR [1] and EFR [2] are used to compress the speech. The devices used to compress speech are called vocoders. The coded speech requires less than 2 bits per sample. This situation is depicted in Figure 1. Between the base stations, the speech is transmitted in an uncoded form (using PCM

5   companding which requires 8 bits per sample).

The terms coded speech and uncoded speech may be described as follows:

Uncoded speech: refers to the digital speech signal samples typically used in telephony; these samples are either in linear 13-bits per sample form or companded form such as the 8-bits per sample μ-law or A-law PCM form; the typical bit-rate is

10   64 kbps.

Coded speech: refers to the compressed speech signal parameters (also referred to as coded parameters) which use a bit rate typically well below 64kbps such as 13 kbps in the case of the GSM FR and 12.2 kbps in the case of GSM EFR; the compression methods are more extensive than the simple PCM companding scheme;

15   examples of compression methods are linear predictive coding, code-excited linear prediction and multi-band excitation coding [4].

The Tandem-Free Operation (TFO) standard [3] will be deployed in GSM digital cellular networks in the near future. The TFO standard applies to mobile-to-mobile calls. Under TFO, the speech signal is conveyed between mobiles in a

20   compressed form after a brief negotiation period. This eliminates tandem voice codecs during mobile-to-mobile calls. The elimination of tandem codecs is known to improve speech quality in the case where the original signal is clean. The key point to note is

that the speech transmission remains coded between the mobile handsets and is depicted in Figure 2.

Under TFO, the transmissions between the handsets and base stations are coded, requiring less than 2 bits per speech sample. However, 8 bits per speech

5   sample are still available for transmission between the base stations. At the base station, the speech is decoded and then A-law companded so that 8 bits per sample are necessary. However, the original coded speech bits are used to replace the 2 least significant bits (LSBs) in each 8-bit A-law companded sample. Once TFO is established between the handsets, the base stations only send the 2 LSBs in each 8-bit

10   sample to their respective handsets and discard the 6 MSBs. Hence vocoder tandeming is avoided. The process is illustrated in Figure 3.

The echo problem and its traditional solution are shown in Figure 4. In wireline networks, echo occurs due to the impedance mismatch at the 4-wire-to-2-wire hybrids. The mismatch results in electrical reflections of a portion of the far-end

15   signal into the near-end signal. Depending on the channel impulse response of the endpath and network delay, the echo can be annoying to the far end listener. The endpath impulse response is estimated using a network echo canceller (EC) and is used to produce an estimate of the echo signal. The estimate is then subtracted from the near-end signal to remove the echo. After EC processing, any residual echo is

20   removed by the non-linear processor (NLP).

In the case of a digital cellular handset, the echo occurs due to the feedback from the speaker (earpiece) to the microphone (mouthpiece). The acoustic feedback

can be significant and the echo can be annoying, particularly in the case of hands-free phones.

Figure 5 shows the feedback path from the speaker to the microphone in a digital cellular handset. The depicted handset does not have echo cancellation

5   implemented in the handset.

Under TFO in GSM networks, if echo cancellation is implemented in the network, a traditional approach requires decoding the coded speech, processing the resulting uncoded speech and then re-encoding it. Such decoding and re-encoding is necessary because traditional echo cancellers can only operate on the uncoded speech

10   signal. This approach is shown in Figure 6. Some of the disadvantages of this approach are as follows.

1. This approach is computationally expensive due to the need for two decoders and an encoder. Typically, encoders are at least an order of magnitude more complex computationally than decoders. Thus, the

15   presence of an encoder, in particular, is a major computational burden.

2. The delay introduced by the decoding and re-encoding processes is undesirable.

3. A vocoder tandem (i.e. two encoder/decoder pairs placed in series) is introduced in this approach, which is known to degrade speech quality due

20   to quantization effects.

In another straightforward approach, comfort noise generation may be used to mask the echo. Comfort noise generation is used for silence suppression or discontinuous transmission purposes (e.g. [5]). It is possible to use such techniques to completely mask the echo whenever echo is detected. However, such techniques suffer from "choppiness" particularly during double-talk conditions, as well as poor and unnatural background transparency.

The proposed techniques are capable of performing echo control (acoustic or linear) directly on the coded speech (i.e. by direct modification of the coded parameters). Low computational complexity and delay are achieved. Tandeming effects are avoided or minimized, resulting in better perceived quality after echo control. Excellent background transparency is also achieved.

Speech compression, which falls under the category of lossy source coding, is commonly referred to as speech coding. Speech coding is performed to minimize the bandwidth necessary for speech transmission. This is especially important in wireless telephony where bandwidth is scarce. In the relatively bandwidth abundant packet networks, speech coding is still important to minimize network delay and jitter. This is because speech communication, unlike data, is highly intolerant of delay. Hence a smaller packet size eases the transmission through a packet network. The four ETSI GSM standards of concern are listed in Table 1.

Table 1: GSM Speech Codecs

Codec Name	Coding Method	Bit Rate (kbits/sec)
Half Rate (HR)	VSELP	5.6
Full Rate (FR)	RPE-LTP	13
Enhanced Full Rate (EFR)	ACELP	12.2
Adaptive Multi-Rate (AMR)	MR-ACELP	5.4-12.2

In speech coding, a set of consecutive digital speech samples is referred to as a speech frame. The GSM coders operate on a frame size of 20ms (160 samples at 8kHz sampling rate). Given a speech frame, a speech encoder determines a small set of

5   parameters for a speech synthesis model. With these speech parameters and the speech synthesis model, a speech frame can be reconstructed that appears and sounds very similar to the original speech frame. The reconstruction is performed by the speech decoder. In the GSM vocoders listed above, the encoding process is much more computationally intensive than the decoding process.

10   The speech parameters determined by the speech encoder depend on the speech synthesis model used. The GSM coders in Table 1 utilize linear predictive coding (LPC) models. A block diagram of a simplified view of a generic LPC speech synthesis model is shown in Figure 7. This model can be used to generate speech-like signals by specifying the model parameters appropriately. In this example speech

15   synthesis model, the parameters include the time-varying filter coefficients, pitch periods, codebook vectors and the gain factors. The synthetic speech is generated as follows. An appropriate codebook vector, $c(n)$, is first scaled by the codebook gain factor G. Here n denotes sample time. The scaled codebook vector is then filtered by a pitch synthesis filter whose parameters include the pitch gain, g_p, and the pitch

20   period, T. The result is sometimes referred to as the total excitation vector, $u(n)$. As implied by its name, the pitch synthesis filter provides the harmonic quality of voiced speech. The total excitation vector is then filtered by the LPC synthesis filter which specifies the broad spectral shape of the speech frame and the broad spectral shape of the corresponding audio signal.

For each speech frame, the parameters are usually updated more than once. For instance, in the GSM FR and EFR coders, the codebook vector, codebook gain and the pitch synthesis filter parameters are determined every subframe (5ms). The LPC synthesis filter parameters are determined twice per frame (every 10ms) in EFR and once per frame in FR.

A typical sequence of steps used in a speech encoder is as follows:

1. Obtain a frame of speech samples.

2. Multiply the frame of samples by a window (e.g. Hamming window) and determine the autocorrelation function up to lag M.

3. Determine the reflection coefficients and/or LPC coefficients from the autocorrelation function. (Note that reflection coefficients are an alternative representation of the LPC filter coefficients.)

4. Transform the reflection coefficients or LPC filter coefficients to a different form suitable for quantization (e.g. log-area ratios or line spectral frequencies)

5. Quantize the transformed LPC coefficients using vector quantization techniques.

6. Add any additional error correction/detection, framing bits etc.

7. Transmit the coded parameters.

The following sequence of operations is typically performed for each subframe by the speech encoder:

1. Determine the pitch period.

2. Determine the corresponding pitch gain.

3. Quantize the pitch period and pitch gain.

4. Inverse filter the original speech signal through the quantized LPC synthesis filter to obtain the LPC residual signal.

5. Inverse filter the LPC residual signal through the pitch synthesis filter to obtain the pitch residual.

6. Determine the best codebook vector.

7. Determine the best codebook gain.

8. Quantize the codebook gain and codebook vector.

9. Update the filter memories appropriately.

A typical sequence of steps used in a speech decoder is as follows:

First, perform any error correction/detection and framing.

Then, for each subframe:

1. Dequantize all the received coded parameters (LPC coefficients, pitch period, pitch gain, codebook vector, codebook gain).

2. Scale the codebook vector by the codebook gain and filter it using the pitch synthesis filter to obtain the LPC excitation signal.

3. Filter the LPC excitation signal using the LPC synthesis filter to obtain a preliminary speech signal.

5   4. Construct a post-filter (usually based on the LPC coefficients).

5. Filter the preliminary speech signal to reduce quantization noise to obtain the final synthesized speech.

As an example of the arrangement of coded parameters in the bit-stream transmitted by the encoder, the GSM FR vocoder is considered. For the GSM FR

10   vocoder, a frame is defined as 160 samples of speech sampled at 8kHz, i.e. a frame is 20ms long. With A-law PCM companding, 160 samples would require 1280 bits for transmission. The encoder compresses the 160 samples into 260 bits. The arrangement of the various coded parameters in the 260 bits of each frame is shown in Figure 8. The first 36 bits of each coded frame consists of the log-area ratios which

15   correspond to LPC synthesis filter. The remaining 224 bits can be grouped into 4 subframes of 56 bits each. Within each subframe, the coded parameter bits contain the pitch synthesis filter related parameters followed by the codebook vector and gain related parameters.

BRIEF SUMMARY OF THE INVENTION

20   The preferred embodiment is useful in a communications system for transmitting a near end digital signal using a compression code comprising a plurality

of parameters including a first parameter. The parameters represent an audio signal comprising a plurality of audio characteristics. The compression code is decodable by a plurality of decoding steps. The communications system also transmits a far end digital signal using a compression code. In such an environment, the echo in the near

5   end digital signal can be reduced by reading at least the first parameter of the plurality of parameters in response to the near end digital signal. At least one of the plurality of the decoding steps is performed on the near end digital signal and the far end digital signal to generate at least partially decoded near end signals and at least partially decoded far end signals. The first parameter is adjusted in response to the at

10   least partially decoded near end signals and at least partially decoded far end signals to generate an adjusted first parameter. The first parameter is replaced with the adjusted first parameter in the near end digital signal. The reading, generating and adjusting preferably are performed by a processor.

Another embodiment of the invention is useful in a communications system

15   for transmitting a near end digital signal comprising code samples further comprising first bits using a compression code and second bits using a linear code. The code samples represent an audio signal having a plurality of audio characteristics. The system also transmits a far end digital signal. In such an environment, any echo in the near end digital signal can be reduced without decoding the compression code by

20   adjusting the first bits and second bits in response to the near end digital signal and the far end digital signal.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic block diagram of a system for speech transmission in a GSM digital cellular network.

Figure 2 is a schematic block diagram of a system for speech transmission in a GSM network under tandem-free operation (TFO).

5 Figure 3 is a graph illustrating transmission of speech under tandem-free operation (TFO).

Figure 4 is a schematic block diagram of a traditional solution to an echo problem in a wireline network.

Figure 5 is a schematic block diagram illustrating acoustic feedback from a 10 speaker to a microphone in a digital cellular telephone.

Figure 6 is a schematic block diagram of a traditional echo cancellation approach for coded speech.

Figure 7 is a schematic block diagram of a generic linear predictive code (LPC) speech synthesis model or speech decoder model.

15 Figure 8 is a diagram illustrating the arrangement of coded parameters in the bit stream for GSM FR.

Figure 9 is a schematic block diagram of a preferred form of coded domain echo control system for acoustic echo environments made in accordance with the invention.

Figure 10 is a schematic block diagram of another preferred form of coded domain echo control system for echo due to 4-wire-to-2-wire hybrids made in accordance with the invention.

Figure 11 is a schematic block diagram of a simplified end path model with flat delay and attenuation.

Figure 12 is a graph illustrating a preliminary echo likelihood versus near end to far end subframe power ratio.

Figure 13 is a flow diagram illustrating a preferred form of coded domain echo control methodology.

Figure 14 is a graph illustrating an exemplary pitch synthesis filter magnitude frequency response.

Figure 15 is a graph illustrating exemplary magnitude frequency responses of an original LPC synthesis filter and flattened versions of such a filter.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The preferred embodiments will be described with reference to the following abbreviations:

ACELP	Algebraic Code Excited Linear Prediction
AE	Audio Enhancer
ALC	Adaptive or Automatic Level Control
CD	Coded Domain or Compressed Domain
CDEC	Coded Domain Echo Control
EFR	Enhanced Full Rate
ETSI	European Telecommunications Standards Institute
FR	Full Rate

GSM	Global System for Mobile Communications
ITU	International Telecommunications Union
MR-ACELP	Multi-Rate ACELP
PCM	Pulse Code Modulation (ITU G.711)
RPE-LTP	Regular Pulse Excitation - Long Term Prediction
TFO	Tandem Free Operation
VSELP	Vector Sum Excitation Linear Prediction

Speech Synthesis Transfer Function

Although many non-linearities and heuristics are involved in the speech synthesis at the decoder, the following approximate transfer function may be attributed to the synthesis process:

$$H(z) = \frac{G}{\left(1 - g_p z^{-T}\right)\left(1 - \sum_{k=1}^{M} a_k z^{-k}\right)} \tag{1}$$

The codebook vector, $c(n)$, is filtered by $H(z)$ to result in the synthesized speech. The key point to note about this generic LPC speech synthesis or decoder model for speech decoding is that the available coded parameters that can be modified to achieve echo control are:

     1. $c(n)$: codebook vector

     2. G : codebook gain

     3. g_p : pitch gain

     4. T : pitch period

5. $\{a_k, k = 1,...,M\}$: LPC coefficients

Most LPC-based vocoders use parameters similar to the above set, parameters that may be converted to the above forms, or parameters that are related to the above forms. For instance, the LPC coefficients in LPC-based vocoders may be represented using log-area ratios (e.g. the GSM FR) or line spectral frequencies (e.g. GSM EFR); both of these forms can be converted to LPC coefficients. An example of a case where a parameter is related to the above form is the block maximum parameter in the GSM FR vocoder; the block maximum can be considered to be directly proportional to the codebook gain in the model described by equation (1).

Thus, although the discussion of coded parameter modification methods is mostly limited to the generic speech decoder model, it is relatively straightforward to tailor these methods for any LPC-based vocoder, and possibly even other models.

It should also be clear that non-linear processing methods such as center-clipping used with uncoded speech for echo control cannot be used on the coded parameters because the coded parameter representation of the speech signal is significantly different. Even the codebook vector signal, $c(n)$, is not amenable to center-clipping due to the significant quantization involved. In many vocoders, the majority of the codebook vector samples are already zero while the non-zero pulses are highly quantized. Hence such non-linear processing approaches are not applicable or effective.

In this specification and claims, the terms linear code and compression code have the following meanings:

Linear code: By a linear code, we mean a compression technique that results in one coded parameter or coded sample for each sample of the audio signal. Examples of linear codes are PCM (A-law and μ-law) ADPCM (adaptive differential pulse code modulation), and delta modulation.

5

Compression code: By a compression code, we mean a technique that results in fewer than one coded parameter for each sample of the audio signal. Typically, compression codes result in a small set of coded parameters for each block or frame of audio signal samples. Examples of compression codes are linear predictive coding

10   based vocoders such as the GSM vocoders (HR, FR, EFR).

Coded Domain Echo Control

Overview

Figure 9 shows a novel implementation of coded domain echo control (CDEC)

15   for a situation where acoustic echo is present. A communications system 10 transmits

near end coded digital signals over a network 24 using a compression code, such as

any of the codes used by the Codecs identified in Table 1. The compression code is

generated by an encoder 16 from linear audio signals generated by a near end

microphone 14 within a near end speaker handset 12. The compression code

20   comprises parameters, such as the shown in Figure 8. The parameters represent an

audio signal comprising a plurality of audio characteristics, including audio level and

power. The compression code is decodable by various decoding steps. As will be

explained, system 10 controls echo in the near end digital signals due to the presence

of a far end digital signals transmitted by system 10 over a network 32. The echo is

25   controlled with minimal delay and minimal, if any, decoding of the compression code

parameters shown in Figure 8.

Near end digital signals using the compression code are received on a near end

terminal 20, and digital signals using an adjusted compression code are transmitted by

a near end terminal 22 over a network 24 to a far end handset (not shown) which includes a decoder (not shown) of the adjusted compression code. Note that the adjusted compression code is compatible with the original compression code. In other words, when the coded parameters are modified or adjusted, we term it the adjusted

5   compression code, but it still is decodable using a standard decoder corresponding to the original compression code. A linear far end audio signal is encoded by a far end encoder (not shown) to generate far end digital signals using a compression code compatible with decoder 18, and is transmitted over a network 32 to a far end terminal 34. A decoder 18 of near end handset 12 decodes the far end digital signals. As

10 shown in Figure 9, echo signals from the far end signals may find their way to encoder 16 of the near end handset 12 through acoustic feedback.

A processor 40 performs various operations on the near end and far end compression code. Processor 40 may be a microprocessor, microcontroller, digital signal processor, or other type of logic unit capable of arithmetic and logical

15 operations.

For each type of codec, a different coded domain echo control algorithm 44 is executed by processor 40 at all times – under compressed mode and linear mode, during TFO as well as non-TFO. A partial decoder 48 is executed by processor 40 to read at least a first of the parameters received at terminal 20. Another partial decoder

20 46 is executed by processor 40 to generate at least partially decoded far end signals. Decoder 48 generates at least partially decoded near end signals. (Note that the compression codes used by the near end and far end signals may be different, and hence the partial decoders may also be different.) Based on the partial decoding,

algorithm 44 generates an echo likelihood signal at least estimating the amount of echo in the near end digital signal. The echo likelihood signal varies over time since the amount of echo depends on the far end speech signal. The echo likelihood signal is used by algorithm 44 to adjust the parameter(s) read by algorithm 44. The adjusted

5   parameter is written into the near end digital signal to form an adjusted near end digital signal which is transmitted from terminal 22 to network 24. In other words, the adjusted parameter is substitued for the originally read parameter. The partial decoders 46 and 48 shown within the Network ALC Device are algorithms executed by processor 40 and are codec-dependent.

10   The partial decoders operate on signals compressed using compression codes. In the case where processor 40 is implemented in a TFO environment, partial decoder 46 may decode the linear code rather than the compression code. Also, in this case, partial decoder 48 decodes the linear code and only determines the coded parameters from the compression code without actually synthesizing the audio signal from the

15   compression code.

Blocks 44, 46 and 48 also may be implemented as hardwired circuits.

Figure 10 shows that the Figure 9 embodiment can be useful for a system in which the echo is due to a 4-wire-to-2-wire hybrid.

The CDEC device/algorithm removes the effects of echo from the near-end

20   coded speech by directly modifying the coded parameters in the bit-stream received from the near-end. Decoding of the near-end and far-end signals is performed in order

to determine the likelihood of echo being present in the near-end. Certain statistics are measured from the decoded signals to determine this likelihood value.

Partial Decoding

The decoding of near-end and far-end signals may be complete or partial depending on the vocoder being used for the encode and decode operations. Some examples of situations where partial decoding suffices are listed below:

1. In code-excited linear prediction (CELP) vocoders, a post-filtering process is performed on the signal decoded using the LPC-based model. This post-filtering process reduces quantization noise. However, since it does not significantly affect the measurement of the statistics necessary for determining the likelihood of echo, the post-filtering stage can be avoided for economy.

2. Under TFO in GSM networks, the CDEC device may be placed between the base station and the switch (known as the A-interface) or between the two switches. Since the 6 MSBs of each 8-bit sample of the speech signal corresponds to the PCM code as shown in Figure 3, it is possible to avoid decoding the coded speech altogether in this situation. A simple table-lookup is sufficient to convert the 8-bit companded samples to 13-bit linear speech samples using A-law companding tables. This provides an economical way to obtain a version of the speech signal without invoking the appropriate decoder. Note that the speech signal obtained in this manner is somewhat noisy,

but has been found to be adequate for the measurement of the statistics

necessary for determining the likelihood of echo.

Determination Of Echo Likelihood

Assuming that some uncoded version (either fully or partially decoded) of the

5   far-end and near-end signals are available, certain statistics are measured and used to

determine the likelihood of echo being present in the near-end signal. The echo

likelihood is estimated for each speech subframe, where the subframe duration is

dependent on the vocoder being used. A preferred approach is described in this

section.

10   A simplified model of the end-path is assumed as shown in Figure 11. The

end-path is assumed to consist of a flat delay of τ samples and an echo return loss

(ERL), λ.

In this model, $s_{NE}(n)$ and $s_{FE}(n)$ are the near-end and far-end uncoded

signals, respectively. It is assumed that the range of τ is known for a given

15   implementation of CDEC, and is specified as follows:

$$\tau_{min} \leq \tau \leq \tau_{max} \tag{2}$$

This assumption is reasonable since the maximum and minimum end-path

delays depend mostly on the speech encoding, speech decoding, channel encoding,

channel decoding and other known transmission delays. The ERL range is assumed to

20   be:

$$0 \leq \lambda \leq 1 \tag{3}$$

The echo likelihood estimation process uses the following variables:

P_{NE} is the power of the current subframe of the near-end signal.

$P_{FE}(0)$ is the power of the current subframe of the far-end signal.

$P_{FE}(m)$ is the power of the mth subframe before the current subframe of the far-end signal. In other words, a buffer of past values of far-end subframe power values is maintained. The buffer size is $B_{max} = \lceil \tau_{max} / N \rceil$ so that the subframe power of the far-end signal up to the maximum possible end-path delay is available. Here N is the number of samples in a subframe.

R is the near-end to far-end subframe power ratio.

ρ_1 is the preliminary echo likelihood.

ρ is the echo likelihood obtained by smoothing the preliminary echo likelihood.

The echo likelihood is estimated for each subframe using the steps below. For some vocoders, particularly lower bit rate vocoders such as GSM HR, the processing may be more appropriately performed frame-by-frame rather than subframe-by-subframe.

Determine the power of $s_{NE}(n)$ for the current subframe as $P_{NE} = \frac{1}{N} \sum_{n=0}^{N-1} s_{NE}^2(n)$.

Determine the power of $s_{FE}(n)$ for the current subframe as

$$P_{FE}(0) = \frac{1}{N}\sum_{n=0}^{N-1} s_{FE}^2(n).$$

Determine the near-end to far-end power ratio as

$$R = \frac{P_{NE}}{\max\left[P_{FE}(B_{min}), P_{FE}(B_{min}+1), ..., P_{FE}(B_{max})\right]}$$ where $B_{min} = \lfloor \tau_{min}/N \rfloor$. The

denominator is essentially the maximum far-end subframe power measured during the

expected end-path delay time period.

· Shift the far-end power values in the buffer, i.e.

$$P_{FE}(B_{max}) = P_{FE}(B_{max}-1); ...; P_{FE}(1) = P_{FE}(0).$$

Determine the preliminary echo likelihood as

$$\rho_1 = \begin{cases} 0 & , \text{ for } R > 63 \\ -0.016R+1.008, & \text{ for } 0.5 \le R \le 63. \\ 1 & , \text{ for } R < 0.5 \end{cases}$$

Smooth the preliminary echo likelihood to obtain the echo likelihood using

$$\rho = 0.9\rho + 0.1\rho_1$$

The graph for the preliminary likelihood as a function of near-end to far-end

subframe power ratio is shown in Figure 12.

Coded Parameter Modification

In this section, the preferred techniques for direct modification of the coded

parameters based on the echo likelihood are described. The direct modification of

each coded parameter of the generic speech decoder model of Figure 7 is first described. Then the corresponding method for modification of the parameters for a standard-based vocoder is described. As an example of a standard-based vocoder, the GSM FR vocoder is considered. After each parameter is modified and quantized according to the standard, the appropriate parameters in the bit-stream are modified appropriately. The preferred embodiment of the overall process is depicted in Figure 13.

Codebook Gain Modification

The codebook gain parameter, G, for each subframe is reduced by a scale factor depending on the echo likelihood, ρ, for the subframe. The modified codebook gain parameter, denoted by G_{new}, is given by:

$$G_{new} = (1 - \rho)G \qquad (4)$$

This parameter is then requantized according to the vocoder standard. Note that the codebook gain controls the overall level of the synthesized signal in the speech decoder model of Figure 7, and therefore controls the overall level of the corresponding audio signal. Attenuating the codebook gain in turn results in the attenuation of the echo.

For the GSM FR, the block maximum parameter, x_{max}, is directly proportional to the codebook gain parameter of the generic model of Figure 7. Hence the modified block maximum parameter is computed as

$$x_{max,new} = (1-\rho)x_{max} \quad (5)$$

$x_{max,new}$ is then requantized according to the method prescribed in the standard.

The resulting 6 bit value is reinserted at the appropriate positions in the bit-stream.

Codebook Vector Modification

5   The codebook vector, $c(n)$, is modified by randomizing the pulse positions

and amplitudes. Randomizing the codebook vector results in destroying the

correlation properties of the echo. This has the effect of destroying much of the

"speech-like" nature of the echo. The randomization is performed whenever the

likelihood of echo is determined to be high, preferably when $\rho > 0.8$. The

10   randomization may be performed using any suitable pseudo-random bit generation

technique.

In the case of the GSM FR, the codebook vector for each subframe is

determined by the RPE grid position parameter (2 bits) and 13 RPE pulses (3 bits

each). These 41 bits are replaced with 41 random bits using a pseudo-random bit

15   generator.

Pitch Synthesis Filter Modification

The pitch synthesis filter implements any period of long-term correlation in

the speech signal, and is particularly important for modeling the harmonics of voiced

speech. The model of this filter discussed in Figure 7 uses only two parameters, the

20   pitch period, T, and the pitch gain, g_p. During voiced speech, the pitch period is

relatively constant over several subframes or frames. The pitch gain in most vocoders ranges from zero to one or a small value above one (e.g. 1.2 in GSM EFR). During strong voiced speech, the pitch gain is at or near its maximum value.

If only echo is present in the near-end signal, the voiced harmonics of the echo

5 are generally well modeled by the pitch synthesis filter; the likelihood of echo is detected to be high ($\rho > 0.8$).

If both echo and near-end speech are present in the near-end signal during a frame period, the likelihood of echo is at moderate levels ($0.5 \leq \rho \leq 0.8$). In such situations, the encoding process generally results in modeling the stronger of the two

10 signals. It is reasonable to assume that, in most cases, the near-end speech is stronger than the echo. If this is the case, then the encoding process, due to its nature, tends to model mostly the near-end speech harmonics and little or none of the echo harmonics with the pitch synthesis filter.

In order to remove or mask voiced echo, the harmonic nature of the echo is

15 destroyed. This is achieved by modifying the pitch synthesis filter parameters as follows:

The pitch period is randomized so that long-term correlation in the echo is removed, hence destroying the voiced nature of the echo. Such randomization is performed only when the likelihood of echo is high, preferably when $\rho > 0.8$.

20 The pitch gain is reduced so as to control the strength of the harmonics or the strength of the long-term correlation in the audio signal. Such gain attenuation is

preferably performed only when the likelihood of echo is at least moderate ($\rho > 0.5$). The new pitch gain is obtained as

$$g_{P,new} = \begin{cases} (1-\rho)g_P & \text{if } \rho > 0.5 \\ g_P & \text{otherwise} \end{cases} \qquad (6)$$

Note that with this approach, the pitch period is not randomized during moderate echo likelihood but the pitch gain may be attenuated so that the voicing quality of the signal is not as strong.

Figure 14 shows the magnitude frequency responses of a pitch synthesis filter with pitch period $T = 41$. The dotted line is the response for a high pitch gain ($g_P = 0.75$) and the solid line illustrates what happens when the pitch gain is attenuated to $g_P = 0.3$. The strength of the harmonics and long-term correlation of an audio signal can be controlled by modifying this parameter in this manner.

In the GSM FR vocoder, the LTP lag parameter of subframe j, denoted by N_j, corresponds to the pitch period T of the model of Figure 7. N_j takes up 7 bits in the bit-stream and can range from 40 to 120, inclusive. Hence when randomizing N_j, it must be replaced with a random number that is also in this range.

The LTP gain parameter of subframe j of the GSM FR vocoder, denoted by b_j, corresponds to the pitch gain g_P of Figure 7. The modified LTP gain parameter is obtained in a manner similar to equation (6) as

$$b_{j,new} = \begin{cases} (1-\rho)b_j & \text{if } \rho > 0.5 \\ b_j & \text{otherwise} \end{cases} \quad (7)$$

LPC Synthesis Filter Modification

In the generic speech decoder model of Figure 7, the LPC synthesis filter transfer function is $1/\left(1 - \sum_{k=1}^{M} a_k z^{-k}\right)$. This filter provides the broad spectral shaping for the synthesized signal. The magnitude frequency response of this filter may be flattened by replacing the coefficients $\{a_k\}$ with $\{\beta^k a_k\}$ with $0 \le \beta \le 1$. β is termed the *spectral morphing factor*. In other words, the modified transfer function is $1/\left(1 - \sum_{k=1}^{M} a_k \beta^k z^{-k}\right)$. Note that when $\beta = 0$, the original LPC synthesis filter is transformed into an all-pass filter, and when $\beta = 1$, the original filter remains unchanged. For all values of β between 0 and 1, the original filter magnitude frequency response experiences some flattening, with greater flattening as $\beta \to 0$. Note that filter stability is maintained in this transformation.

The effect of such spectral morphing on echo is to reduce or remove any formant structure present in the signal. The echo is blended or morphed to sound like background noise. As an example, the LPC synthesis filter magnitude frequency response for a voiced speech segment and its flattened versions for several different values of β are shown in Figure 15.

In the preferred implementation, the spectral morphing factor β is determined as

$$\beta = \begin{cases} 1 - 0.8\rho & \text{for } \rho > 0.5 \\ 1 & \text{otherwise} \end{cases} \qquad (8)$$

A similar spectral morphing method is obtained for other representations of the LPC filter coefficients commonly used in vocoders such as reflection coefficients, log-area ratios, inverse sines functions, and line spectral frequencies.

5   For example, the GSM FR vocoder utilizes log-area ratios for representing the LPC synthesis filter. Given the 8 log-area ratios corresponding to a frame, denoted by $LAR(i)$, $i = 1, 2, ...8$, the spectrally morphed log-area ratios are obtained using

$$LAR_{new}(i) = \beta LAR(i) \qquad (9)$$

where β is determined according to equation (8). This method spectrally

10   flattens the LPC synthesis filter magnitude frequency response. Alternatively, in order to morph the log-area ratios towards a predetermined spectrum or magnitude frequency response, such as the background noise spectrum represented by a set of log-area ratios denoted by $LAR_{noise}(i)$, the appropriate morphing equation is

$$LAR_{new}(i) = \beta LAR(i) + (1 - \beta) LAR_{noise}(i) \qquad (10)$$

15   The modified log-area ratios are then quantized according to the specifications in the standard. Note that these approaches to modification of the log-area ratios preserve the stability of the LPC synthesis filter.

An exemplary approach for background noise spectrum estimation and representation of filter coefficients comprising log-area ratios corresponding to the

vocoder and an LPC filter is provided in the comfort noise generation standard [5] and the references therein.

When line spectral frequencies are used for representing the LPC synthesis filter (e.g. the GSM EFR), an approach similar to that for log-area ratios is also appropriate. Denote the line spectral frequencies by $f_i, i = 1,..M$, where M is the order of the LPC synthesis filter which is assumed even (typical). When the line spectral frequencies are evenly spaced apart from 0 to half the sampling frequency, the resulting LPC synthesis filter will be all-pass (i.e. flat magnitude frequency response). Denote the set of line spectral frequencies corresponding to such a spectrally flat LPC filter by $f_{i,flat}, i = 1,..,M$. Then, the spectrally morphed line spectral frequencies are obtained using

$$f_{i,new} = \beta f_i + (1 - \beta) f_{i,flat} \qquad (11)$$

where β is determined according to equation (8). This method spectrally flattens the LPC synthesis filter magnitude frerquency response. Alternatively, in order to morph the line spectral frequencies towards a predetermined spectrum or magnitude frequency response, such as the background noise spectrum represented by a set of line spectral frequencies denoted by $f_{i,noise}$, the appropriate morphing equation is

$$f_{i,new} = \beta f_i + (1 - \beta) f_{i,noise} \qquad (12)$$

The modified line spectral frequencies are then quantized according to the specifications in the standard. Note that these approaches to modification of the line spectral frequencies preserve the stability of the LPC synthesis filter. Appropriate techniques for background noise spectrum estimation and representation of filter

5   coefficients comprising line spectral frequencies may be found in the corresponding vocoder standards on comfort noise generation.

Minimal Delay Technique

Large buffering, processing and transmission delays are already present in cellular networks without any network voice quality enhancement processing. Further

10   network processing of the coded speech for speech enhancement purposes will add additional delay. Minimizing this delay is important to speech quality. In this section, a novel approach for minimizing the delay is discussed. The example used is the GSM FR vocoder.

Figure 8 shows the order in which the coded parameters from the GSM FR

15   encoder are received. A straightforward approach involves buffering up the entire 260 bits for each frame and then processing these buffered bits for coded domain echo control purposes. However, this introduces a buffering delay of about 20ms plus the processing delay.

It is possible to minimize the buffering delay as follows. First, note that the

20   entire first subframe can be decoded as soon as bit 92 is received. Hence the first subframe may be processed after about 7.1ms (20ms times 92/260) of buffering delay. Hence the buffering delay is reduced by almost 13ms.

When using this novel low delay approach, the coded LPC synthesis filter parameters are modified based on information available at the end of the first subframe of the frame. In other words, the entire frame is affected by the echo likelihood computed based on the first subframe. In experiments conducted, no

5   noticeable artifacts were found due to this 'early' decision, particularly because the echo likelihood is a smoothed quantity based effectively on several previous subframes as well as the current subframe.

Update of Error Correction/Detection Bits and Framing Bits

When applying the novel coded domain processing techiques described in this

10 report for removing echo, some are all of the bits corresponding to the coded parameters are modified in the bit-stream. This may affect other error-correction or detection bits that may also be embedded in the bit-stream. For instance, a speech encoder may embed some checksums in the bit-stream for the decoder to verify to ensure that an error-free frame is received. Such checksums as well as any parity

15 check bits, error correction or detection bits, and framing bits are updated in accordance with the appropriate standard, if necessary.

Operation under the GSM Tandem Free Operation Standard

If only the coded parameters are available, then partial or full decoding may be performed as explained earlier, whereby the coded parameters are used to reconstruct

20 a version of the audio signal. However, when operating under a situation such as the GSM TFO environment, additional information is available in addition to the coded parameters. This additional information is the 6 MSBs of the A-law PCM samples of

the audio signal. In this case, these PCM samples may be used to reconstruct a version of the audio signal for both the far end and near end without using the coded parameters. This results in computational savings.

Those skilled in the art of communications will recognize that the preferred embodiments can be modified and altered without departing from the true spirit and scope of the invention as defined in the appended claims.

What is claimed is:

1.   In a communications system for transmitting a near end digital signal using a compression code comprising a plurality of parameters including a first parameter, said parameters representing an audio signal comprising a plurality of audio characteristics,

5   said compression code being decodable by a plurality of decoding steps, said communications system also transmitting a far end digital signal using a compression code, apparatus for reducing echo in said near end digital signal comprising:

a processor responsive to said near end digital signal to read at least said first parameter of said plurality of parameters, to perform at least one of said plurality of

10   decoding steps on said near end digital signal and said far end digital signal to generate at least partially decoded near end signals and at least partially decoded far end signals, responsive to said at least partially decoded near end signals and at least partially decoded far end signals to adjust said first parameter to generate an adjusted first parameter and to replace said first parameter with said adjusted first parameter in said

15   near end digital signal.

2.   Apparatus, as claimed in claim 1, wherein said first parameter is a quantized first parameter and wherein said processor generates said adjusted first parameter in part by quantizing said adjusted first parameter before writing said adjusted first parameter into said near end digital signal.

20   3.   Apparatus, as claimed in claim 1, wherein said processor is responsive to said at least partially decoded near end signals and said at least partially decoded far end signals to generate an echo likelihood signal representing the amount of echo present in

said partially decoded near end signals, and wherein said processor is responsive to said echo likelihood signal to adjust said first parameter.

4.   Apparatus, as claimed in claim 3, wherein said characteristics comprise spectral shape and wherein said first parameter comprises a representation of filter coefficients, and wherein said processor is responsive to said echo likelihood signal to adjust said representation of filter coefficients towards a magnitude frequency response.

5.   Apparatus, as claimed in claim 4, wherein said representation of filter coefficients comprises line spectral frequencies.

6.   Apparatus, as claimed in claim 4, wherein said representation of filter coefficients comprises log area ratios.

7.   Apparatus, as claimed in claim 4, wherein said magnitude frequency response corresponds to background noise.

8.   Apparatus, as claimed in claim 1, wherein said characteristics comprise the overall level of said audio signal and wherein said first parameter comprises codebook gain.

9.   Apparatus, as claimed in claim 1, wherein said first parameter comprises a codebook vector parameter.

10.   Apparatus, as claimed in claim 1, wherein said characteristics comprise period of long-term correlation and wherein said first parameter comprises a pitch period parameter.

11.   Apparatus, as claimed in claim 1, wherein said characteristics comprise strength of long-term correlation and wherein said first parameter comprises a pitch gain parameter.

12.   Apparatus, as claimed in claim 1, wherein said characteristics comprise spectral shape and wherein said first parameter comprises a representation of filter coefficients.

13.   Apparatus, as claimed in claim 12, wherein said representation of filter

5   coefficients comprises log area ratios.

14.   Apparatus, as claimed in claim 12, wherein said representation of filter coefficients comprises line spectral frequencies.

15.   Apparatus, as claimed in claim 12, wherein said representation of filter coefficients corresponds to a linear predictive coding synthesis filter.

10   16.   Apparatus, as claimed in claim 1, wherein said first parameter corresponds to a first characteristic of said plurality of audio characteristics, wherein said plurality of decoding steps comprises at least one decoding step avoiding substantial altering of said first characteristic and wherein said processor avoids performing said at least one decoding step.

15   17.   Apparatus, as claimed in claim 16, wherein said audio characteristic comprises power and wherein said first characteristic comprises power.

18.   Apparatus, as claimed in claim 16, wherein said at least one decoding step comprises post-filtering.

19.   Apparatus, as claimed in claim 1, wherein said compression code

20   comprises a linear predictive code.

20.   Apparatus, as claimed in claim 1, wherein said compression code comprises regular pulse excitation – long term prediction code.

21.   Apparatus, as claimed in claim 1, wherein said compression code comprises code-excited linear prediction code.

22.   Apparatus, as claimed in claim 1, wherein said first parameter comprises a

series of first parameters received over time, wherein said processor is responsive to said

near end digital signal to read said series of first parameters, and wherein said processor

is responsive to said at least partially decoded near end and far end signals and to at least

5   a plurality of said series of first parameters to generate said adjusted first parameter.

23.   Apparatus, as claimed in claim 1, wherein said compression code is

arranged in frames of said digital signals and wherein said frames comprise a plurality of

subframes each comprising said first parameter, wherein said processor is responsive to

said compression code to read at least said first parameter from each of said plurality of

10   subframes, and wherein said processor replaces said first parameter with said adjusted

first parameter in each of said plurality of subframes.

24.   Apparatus, as claimed in claim 23, wherein said processor reads said first

parameter from a first of said subframes, begins to perform at least a plurality of said

decoding steps on said near end digital signal during said first subframe and replaces

15   said first parameter with said adjusted first parameter before processing a subframe

following the first subframe so as to achieve lower delay.

25.   Apparatus, as claimed in claim 1, wherein said compression code is

arranged in frames of said digital signals and wherein said frames comprise a plurality of

subframes each comprising said first parameter, wherein said processor performs at least

20   a plurality of said decoding steps during a first of said subframes to generate said at least

partially decoded near end and far end signals, reads said first parameter from a second

of said subframes occurring subsequent to said first subframe, generates said adjusted

first parameter in response to said at least partially decoded near end and far end signals

and said first parameter, and replaces said first parameter of said second subframe with said adjusted first parameter.

26. In a communications system for transmitting a near end digital signal comprising code samples, said code samples comprising first bits using an compression code and second bits using a linear code, said code samples representing an audio signal, said audio signal having a plurality of audio characteristics, said system also transmitting a far end digital signal, apparatus for reducing echo in said near end digital signal without decoding said compression code comprising:

27. a processor responsive to said near end digital signal and said far end digital signal to adjust said first bits and said second bits.

28. Apparatus, as claimed in claim 26, wherein said linear code comprises pulse code modulation (PCM) code.

29. Apparatus, as claimed in claim 26, wherein said compression code samples conform to the tandem-free operation of the global system for mobile communications standard.

30. Apparatus, as claimed in claim 26, wherein said first bits comprise the two least significant bits of said samples and wherein said second bits comprise the 6 most significant bits of said samples.

31. Apparatus, as claimed in claim 29, wherein said 6 most significant bits comprise PCM code.

32. In a communications system for transmitting a near end digital signal using a compression code comprising a plurality of parameters including a first parameter, said parameters representing an audio signal comprising a plurality of audio characteristics, said compression code being decodable by a plurality of decoding steps, said

communications system also transmitting a far end digital signal using a compression code, a method of reducing echo in said near end digital signal comprising:

reading at least said first parameter of said plurality of parameters in response to said near end digital signal;

5   performing at least one of said plurality of decoding steps on said near end digital signal and said far end digital signal to generate at least partially decoded near end signals and at least partially decoded far end signals;

adjusting said first parameter in response to said at least partially decoded near end signals and at least partially decoded far end signals to generate an adjusted first

10   parameter; and

replacing said first parameter with said adjusted first parameter in said near end digital signal.

33.   A method, as claimed in claim 31, wherein said first parameter is a quantized first parameter and wherein said adjusting comprises generating said adjusted

15   first parameter in part by quantizing said adjusted first parameter.

34.   A method, as claimed in claim 31, wherein said adjusting comprises generating an echo likelihood signal representing the amount of echo present in said partially decoded near end signals in response to said at least partially decoded near end signals and said at least partially decoded far end signals, and wherein said adjusting

20   further comprises adjusting said first parameter in response to said echo likelihood signal.

35.   A method, as claimed in claim 33, wherein said characteristics comprise spectral shape and wherein said first parameter comprises a representation of filter coefficients, and wherein said adjusting comprises adjusting said representation of filter

coefficients towards a magnitude frequency response in response to said echo likelihood signal.

36.   A method, as claimed in claim 34, wherein said representation of filter coefficients comprises line spectral frequencies.

37.   A method, as claimed in claim 34, wherein said representation of filter coefficients comprises log area ratios.

38.   A method, as claimed in claim 34, wherein said magnitude frequency response corresponds to background noise.

39.   A method, as claimed in claim 31, wherein said characteristics comprise the overall level of said audio signal and wherein said first parameter comprises codebook gain.

40.   A method, as claimed in claim 31, wherein said first parameter comprises a codebook vector parameter.

41.   A method, as claimed in claim 31, wherein said characteristics comprise period of long-term correlation and wherein said first parameter comprises a pitch period parameter.

42.   A method, as claimed in claim 31, wherein said characteristics comprise strength of long-term correlation and wherein said first parameter comprises a pitch gain parameter.

43.   A method, as claimed in claim 31, wherein said characteristics comprise spectral shape and wherein said first parameter comprises a representation of filter coefficients.

44.   A method, as claimed in claim 42, wherein said representation of filter coefficients comprises log area ratios.

45.   A method, as claimed in claim 42, wherein said representation of filter coefficients comprises line spectral frequencies.

46.   A method, as claimed in claim 42, wherein said representation of filter coefficients corresponds to a linear predictive coding synthesis filter.

5   47.   A method, as claimed in claim 31, wherein said first parameter corresponds to a first characteristic of said plurality of audio characteristics, wherein said plurality of decoding steps comprises at least one decoding step avoiding substantial altering of said first characteristic and wherein said performing at least a plurality of said decoding steps comprises avoiding performing said at least one decoding step.

10   48.   A method, as claimed in claim 46, wherein said audio characteristic comprises power and wherein said first characteristic comprises power.

49.   A method, as claimed in claim 46, wherein said at least one decoding step comprises post-filtering.

50.   A method, as claimed in claim 31, wherein said compression code

15   comprises a linear predictive code.

51.   A method, as claimed in claim 31, wherein said compression code comprises regular pulse excitation – long term prediction code.

52.   A method, as claimed in claim 31, wherein said compression code comprises code-excited linear prediction code.

20   53.   A method, as claimed in claim 31, wherein said first parameter comprises a series of first parameters received over time, wherein said reading comprises reading said series of first parameters, and wherein said adjusting comprises generating said adjusted first parameter in response to said at least partially decoded near end and far end signals and to at least a plurality of said series of first parameters.

54.   A method, as claimed in claim 31, wherein said compression code is
arranged in frames of said digital signals and wherein said frames comprise a plurality of
subframes each comprising said first parameter, wherein said reading comprises reading
at least said first parameter from each of said plurality of subframes in response to said
5   compression code, and wherein said replacing comprises replacing said first parameter
with said adjusted first parameter in each of said plurality of subframes.

55.   A method, as claimed in claim 53, wherein said reading comprises reading
said first parameter from a first of said subframes, wherein said performing comprises
beginning to perform at least a plurality of said decoding steps on said near end digital
10   signal during said first subframe and wherein said replacing comprises replacing said
first parameter with said adjusted first parameter before processing a subframe following
the first subframe so as to achieve lower delay.

56.   A method, as claimed in claim 31, wherein said compression code is
arranged in frames of said digital signals and wherein said frames comprise a plurality of
15   subframes each comprising said first parameter, wherein said performing comprises
performing at least a plurality of said decoding steps during a first of said subframes to
generate said at least partially decoded near end and far end signals, wherein said reading
comprises reading said first parameter from a second of said subframes occurring
subsequent to said first subframe, wherein said adjusting comprises generating said
20   adjusted first parameter in response to said at least partially decoded near end and far end
signals and said first parameter, and wherein said replacing comprises replacing said first
parameter of said second subframe with said adjusted first parameter.

57.   In a communications system for transmitting a near end digital signal
comprising code samples, said code samples comprising first bits using a compression

code and second bits using a linear code, said code samples representing an audio signal, said audio signal having a plurality of audio characteristics, said system also transmitting a far end digital signal, a method of reducing echo in said near end digital signal without decoding said compression code comprising:

5   adjusting said first bits and said second bits in response to said near end digital signal and said far end digital signal.
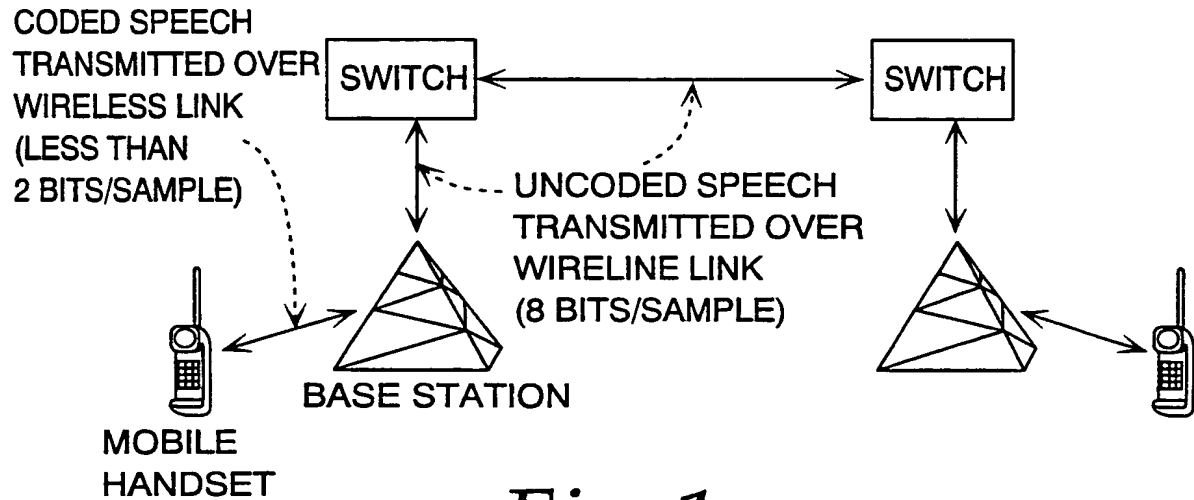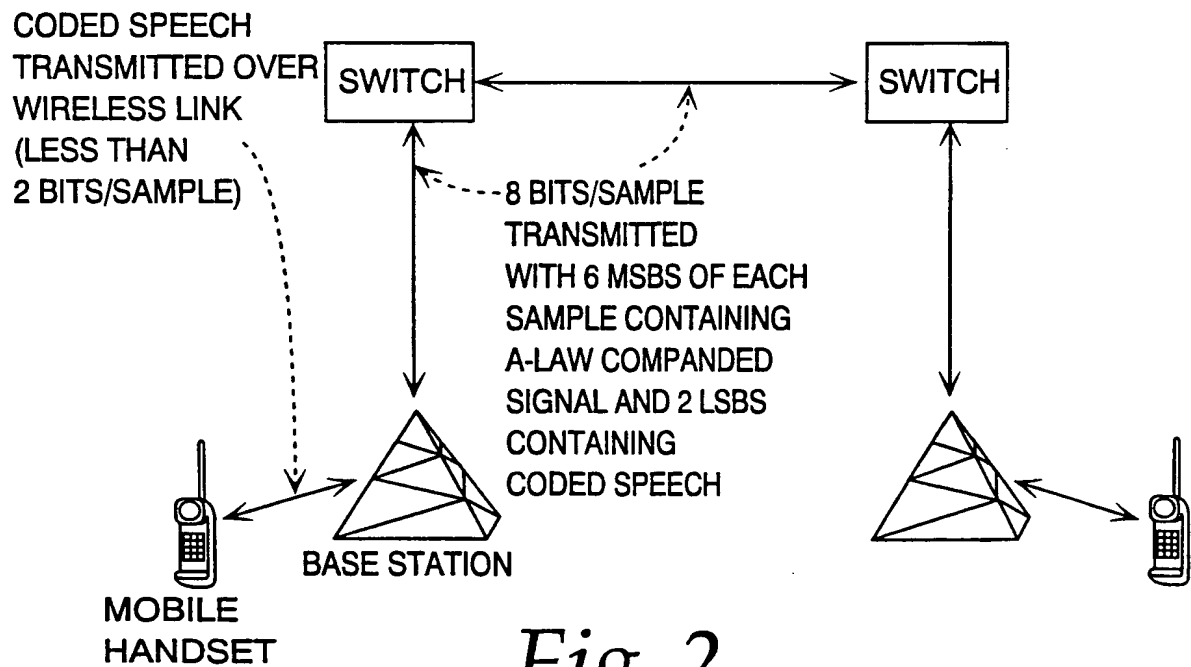
58.   A method, as claimed in claim 56, wherein said linear code comprises pulse code modulation (PCM) code.

59.   A method, as claimed in claim 56, wherein said compression code samples

10 conform to the tandem-free operation of the global system for mobile communications standard.

60.   A method, as claimed in claim 56, wherein said first bits comprise the two least significant bits of said samples and wherein said second bits comprise the 6 most significant bits of said samples.

15   61.   A method, as claimed in claim 59, wherein said 6 most significant bits comprise PCM code.

CODED SPEECH
TRANSMITTED OVER
WIRELESS LINK
(LESS THAN
2 BITS/SAMPLE)

SWITCH ⟷ SWITCH

UNCODED SPEECH
TRANSMITTED OVER
WIRELINE LINK
(8 BITS/SAMPLE)

BASE STATION

MOBILE
HANDSET

Fig. 1

CODED SPEECH
TRANSMITTED OVER
WIRELESS LINK
(LESS THAN
2 BITS/SAMPLE)

SWITCH ⟷ SWITCH

8 BITS/SAMPLE
TRANSMITTED
WITH 6 MSBS OF EACH
SAMPLE CONTAINING
A-LAW COMPANDED
SIGNAL AND 2 LSBS
CONTAINING
CODED SPEECH

BASE STATION

MOBILE
HANDSET

Fig. 2

Fig. 3

2 BITS PER SAMPLE RECEIVED AT DESTINATION MOBILE STATION

TRANSMIT 2 LSBS TO DESTINATION HANDSET

TRANSMIT 8 BITS PER SAMPLE TO DESTINATION BASE STATION

DISCARD 2 LSBS AND REPLACE WITH ORIGINAL CODED SPEECH BITS.

DECODE SPEECH AT RECEIVING BASE STATION

2 BITS PER SAMPLE FROM MOBILE STATION

Fig. 4

NON-LINEAR PROCESSOR (NLP) E.G. CENTER-CLIPPER

ADAPTIVE FILTER

NEAR-END SIGNAL

4-WIRE TO 2-WIRE

FAR-END SIGNAL

ECHO PATH OR ENDPATH

Fig. 5

Fig. 6

Fig. 7

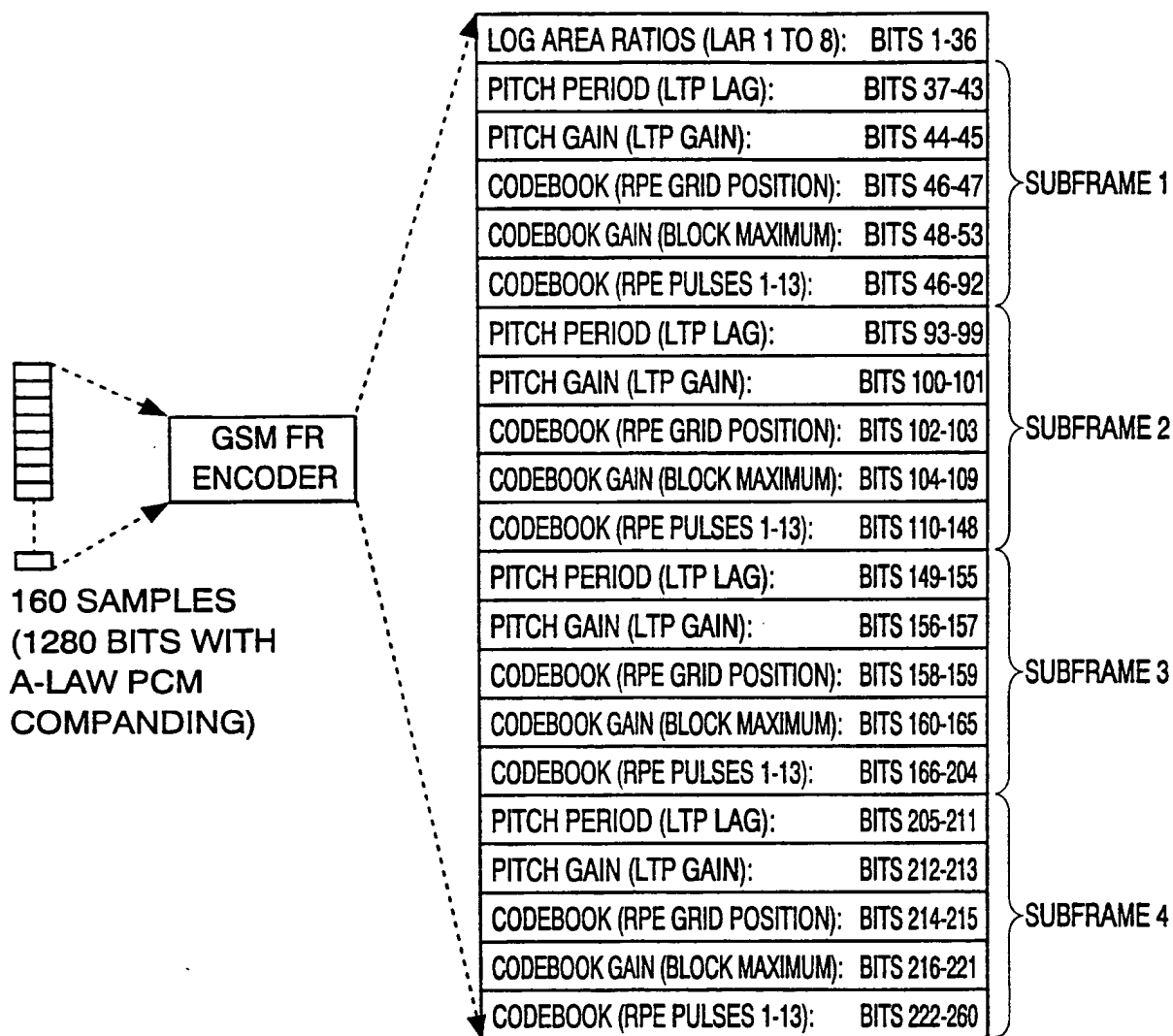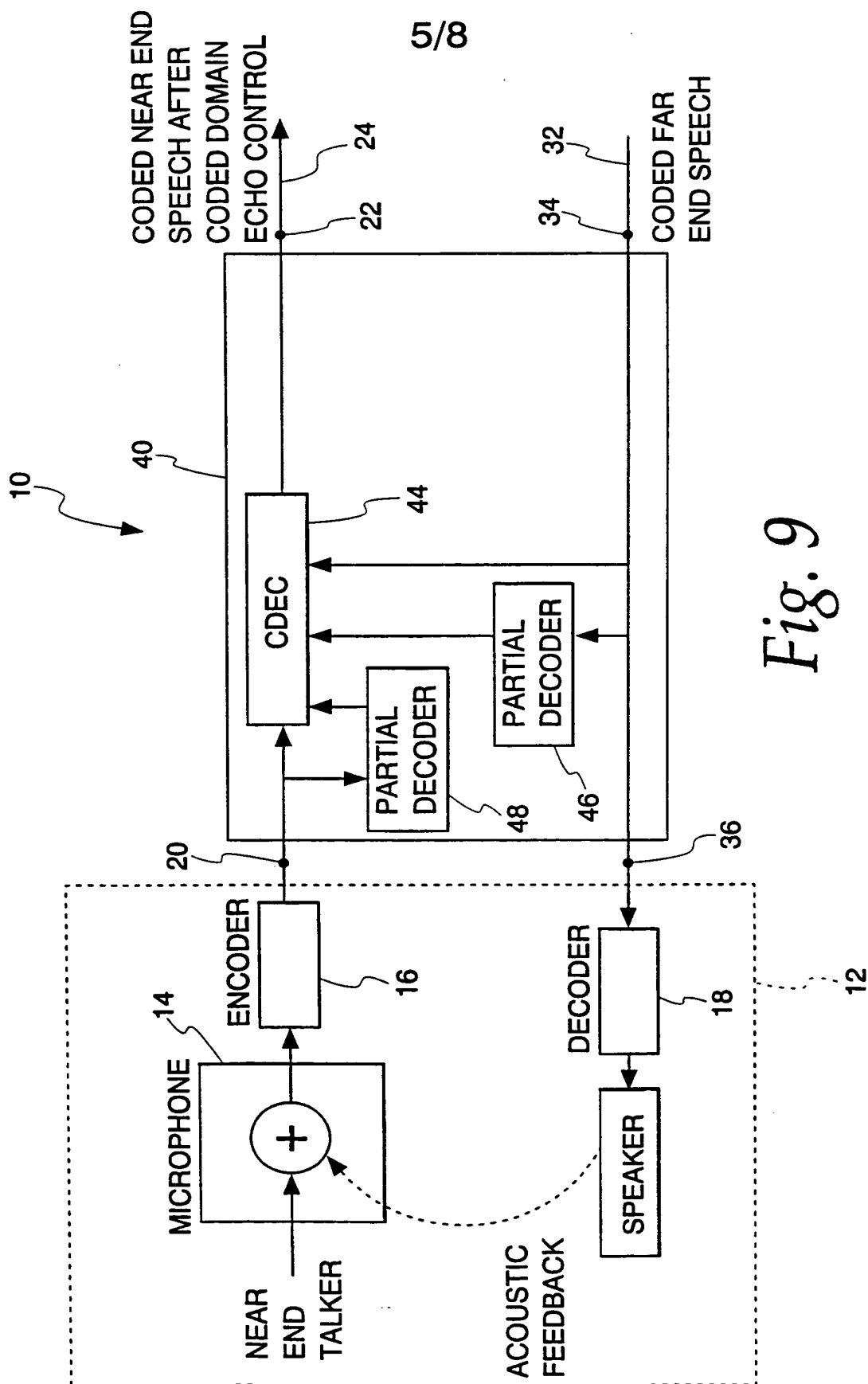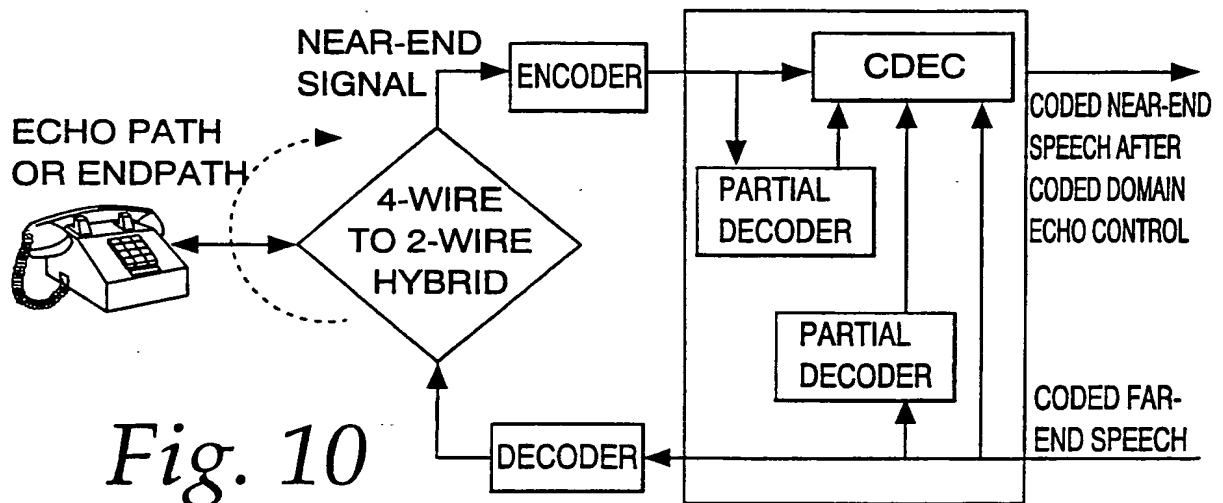| LOG AREA RATIOS (LAR 1 TO 8):	BITS 1-36	
PITCH PERIOD (LTP LAG):	BITS 37-43	SUBFRAME 1
PITCH GAIN (LTP GAIN):	BITS 44-45	
CODEBOOK (RPE GRID POSITION):	BITS 46-47	
CODEBOOK GAIN (BLOCK MAXIMUM):	BITS 48-53	
CODEBOOK (RPE PULSES 1-13):	BITS 46-92	
PITCH PERIOD (LTP LAG):	BITS 93-99	SUBFRAME 2
PITCH GAIN (LTP GAIN):	BITS 100-101	
CODEBOOK (RPE GRID POSITION):	BITS 102-103	
CODEBOOK GAIN (BLOCK MAXIMUM):	BITS 104-109	
CODEBOOK (RPE PULSES 1-13):	BITS 110-148	
PITCH PERIOD (LTP LAG):	BITS 149-155	SUBFRAME 3
PITCH GAIN (LTP GAIN):	BITS 156-157	
CODEBOOK (RPE GRID POSITION):	BITS 158-159	
CODEBOOK GAIN (BLOCK MAXIMUM):	BITS 160-165	
CODEBOOK (RPE PULSES 1-13):	BITS 166-204	
PITCH PERIOD (LTP LAG):	BITS 205-211	SUBFRAME 4
PITCH GAIN (LTP GAIN):	BITS 212-213	
CODEBOOK (RPE GRID POSITION):	BITS 214-215	
CODEBOOK GAIN (BLOCK MAXIMUM):	BITS 216-221	
CODEBOOK (RPE PULSES 1-13):	BITS 222-260	

GSM FR ENCODER

160 SAMPLES
(1280 BITS WITH
A-LAW PCM
COMPANDING)

Fig. 8

Fig. 9

Fig. 10

Fig. 11

Fig. 12

7/8

LPC SYSTHESIS
FILTER PARAMETERS

PITCH SYSTHESIS
FILTER PARAMETERS

CODEBOOK VECTORS
AND GAIN
PARAMETERS

EXTRACT CODED
PARAMETERS
FROM NEAR-END
BIT-STREAM

PARTIALLY
DECODE

EXTRACT CODED
PARAMETERS
FROM FAR-END
BIT-STREAM

PARTIALLY
DECODE

DETERMINE
ECHO
LIKELIHOOD

CODED
PARAMETER
MODIFICATION
ALGORITHMS

INSERT
MODIFIED BITS
BACK INTO BIT-
STREAM

UPDATE ANY
FRAMING OR ERROR
DETECTION AND
CORRECTION BITS
SUCH AS CHECKSUMS
AND PARITY BITS.

QUANTIZATION
OF PARAMETERS
ACCORDING TO
VOCODER
STANDARD

Fig. 13

$T=41, g_P=0.75$

$T=41, g_P=0.3$

Fig. 14

Fig. 15